Preschool Psychopathology Reported by Parents in 23 Societies: Testing the Seven-Syndrome Model of the Child Behavior Checklist for Ages 1.5–5

Masha Y. Ivanova, Ph.D., Thomas M. Achenbach, Ph.D., Leslie A. Rescorla, Ph.D., Valerie S. Harder, Ph.D., Rebecca P. Ang, Ph.D., Niels Bilenberg, M.D., Ph.D., Gudrun Bjarnadottir, Ph.D., Christiane Capron, Ph.D., Sarah S.W. De Pauw, M.A., Pedro Dias, Ph.D., Anca Dobrean, Ph.D., Manfred Doepfner, Ph.D., Michele Duyme, Ph.D., Valsamma Eapen, Ph.D., Nese Erol, Ph.D., Elaheh Mohammad Esmaeili, Ph.D., Lourdes Ezpeleta, Ph.D., Alessandra Frigerio, Ph.D., Miguel M. Gonçalves, Ph.D., Halldor S. Gudmundsson, M.A., Suh-Fang Jeng, Sc.D., Pranvera Jetishi, B.A., Roma Jusiene, Ph.D., Young-Ah Kim, Ph.D., Solvejg Kristensen, M.H.Sc., Felipe Lecannelier, M.A., Patrick W.L. Leung, Ph.D., Jianghong Liu, Ph.D., Rosario Montirosso, M.Sc., Kyung Ja Oh, Ph.D., Julia Plueck, Ph.D., Rolando Pomalima, M.D., Mimoza Shahini, M.D., Jaime R. Silva, Ph.D., Zynep Simsek, Ph.D., Andre Sourander, M.D., Jose Valverde, M.D., Karla G. Van Leeuwen, Ph.D., Bernardine S.C. Woo, M.Med., Yen-Tzu Wu, M.S., Stephen R. Zubrick, Ph.D., Frank C. Verhulst, M.D., Ph.D.

Objective: To test the fit of a seven-syndrome model to ratings of preschoolers' problems by parents in very diverse societies. **Method:** Parents of 19,106 children 18 to 71 months of age from 23 societies in Asia, Australasia, Europe, the Middle East, and South America completed the Child Behavior Checklist for Ages 1.5–5 (CBCL/1.5–5). Confirmatory factor analyses were used to test the seven-syndrome model separately for each society. **Results:** The primary model fit index, the root mean square error of approximation (RMSEA), indicated acceptable to good fit for each society. Although a six-syndrome model combining the Emotionally Reactive and Anxious/Depressed syndromes also fit the data for nine societies, it fit less well than the seven-syndrome model for seven of the nine societies. Other fit indices yielded less consistent results than the RMSEA. **Conclusions:** The seven-syndrome model provides one way to capture patterns of children's problems that are manifested in ratings by parents from many societies. Clinicians working with preschoolers from these societies can thus assess and describe parents' ratings of behavioral, emotional, and social problems in terms of the seven syndromes. The results illustrate possibilities for culture–general taxonomic constructs of preschool psychopathology. Problems not captured by the CBCL/1.5–5 may form additional syndromes, and other syndrome models may also fit the data. J. Am. Acad. Child Adolesc. Psychiatry, 2010;49(12):1215–1224. **Key Words:** preschoolers, assessment, taxonomy, multicultural, confirmatory factor analysis

Preschool children can experience significant and persistent mental health problems.[1] Yet, as Egger and Angold[2] pointed out, "There is currently no consensus about the best criteria for defining most psychiatric disorders in very young children" (p. 313). Reliable and valid taxonomies are needed to guide the growing efforts to understand and treat preschool psychopathology. In reviewing the reliability, validity, and usefulness of five contemporary classification systems for early childhood disorders, Postert et al.[3] concluded that "they must still be seen

This article is discussed in an editorial by Dr. Alice Carter on page 1181.

as limited" (p. 37) and that "the *DSM-IV* and ICD 10 at present offer no comprehensive and recognized classification for diagnosing psychiatric and developmental disorders in preschool children" (p. 30).

One way to develop taxonomies is by statistically identifying patterns of co-occurring problems reported for large samples of children. Descriptively designated as "syndromes," such patterns can provide a basis for taxonomic constructs. Syndromes are often derived via exploratory factor analyses (EFA) and are then tested via confirmatory factor analyses (CFA) of new data sets.

The Child Behavior Checklist for Ages 1.5–5[4] obtains parents' ratings of 99 emotional, behavioral, and social problems, plus one open-ended problem item. Using a combination of EFA and CFA, Achenbach and Rescorla[4] derived seven syndromes from CBCL/1.5–5 ratings of 1,728 preschoolers who were primarily from the United States. Sixty-seven items loaded on the syndromes, which were designated as Emotionally Reactive, Anxious/Depressed, Somatic Complaints, Withdrawn, Sleep Problems, Attention Problems, and Aggressive Behavior. Based on second-order factor analyses, the first four syndromes were found to form a grouping that was designated as "Internalizing," while the last two were found to form a grouping that was designated as "Externalizing." The present study was designed to test the generalizability of the seven-syndrome model in 23 societies.

The generalizability of an assessment instrument's syndrome structure across populations is termed configural invariance.[5] Configural invariance is a component of measurement invariance, which is the degree to which an instrument measures the same constructs in different populations. Although there are additional components of measurement invariance, configural invariance is the most fundamental component. To put invariance in more applied terms, when an instrument that was developed in one society is applied in another society, it is important to demonstrate that it measures the same constructs in the new society. Configural invariance speaks to whether the same items load on the same factors (syndrome constructs) in the new society. Other components of measurement invariance refer to other measurement properties of the instrument, such as whether the strength of item loadings is the same in the two societies (i.e., item loading or "factorial" invariance), or whether individuals in the two societies who have the same scores on the syndrome construct(s) also receive the same item scores (i.e., "intercept invariance").

Previous EFA and CFA Studies of the Preschool CBCL

Konold et al.[6] tested the CBCL/1.5–5 syndrome model using CBCL data obtained in the National Institute of Child Health and Development Study of Early Child Care (NICHD SECC). Mothers of 1,097 children completed an early version of the CBCL/1.5–5, namely the CBCL for Ages 2–3[7], when the children were 24 months old. The study tested whether the CBCL/1.5–5 syndrome model was invariant across gender, ethnic groups (African American vs. white), and socioeconomic status (SES; low vs. high). First, the authors compared the fit of a single factor model, a correlated three-factor model comprising the Internalizing, Externalizing and Sleep Problems factors, and a correlated seven-factor model comprising the CBCL/1.5–5 syndromes. The single- and three-factor models fit the data poorly. The seven-factor model fit the data well when it was reduced to six factors by combining the Emotionally Reactive and Anxious/Depressed factors. Second, multigroup CFAs were performed to test whether the six-factor model had factorial invariance across gender, ethnic groups, and SES. The model was found to have factorial invariance across gender for all factors, except Somatic Complaints. The Sleep Problems and Attention Problems factors were invariant across ethnic groups and SES, whereas the Aggressive Behavior factor was invariant only across ethnic groups.

The Konold et al.[6] results are difficult to interpret for the following reasons: (a) CFAs were applied to subsets of CBCL items ("parcels") formed "by grouping items that were nearest to one another with respect to the magnitude of their factor loadings reported in the manual for the CBCL/1.5–5" (p. 116); (b) the small sample of African Americans (N = 116) casts doubts on the findings for ethnic groups; and (c) because the authors used CBCL/2–3 data, they lacked data for two of the nine items comprising the Emotionally Reactive syndrome, which probably weakened this factor and their test of the seven-syndrome model.

To test the fit of the CBCL/1.5–5 syndrome model to ratings of 707 Chinese girls adopted by U.S. and Canadian parents, Tan et al.[8] replicated the procedures of Achenbach and Rescorla[4] by performing CFAs on tetrachoric correlations. They also performed CFAs on polychoric correlations. CFAs of both types of correlations supported the seven-syndrome model.

Before publication of the seven-syndrome CBCL/1.5–5 model, Koot et al.[9] tested the factor structure of a Dutch translation of the CBCL/2–3.[7] Principal factor analyses were applied separately to Dutch clinical (N = 426), general population (N = 420), and twin (N = 1,306 pairs) samples, and a factor solution that was consistent across the samples was selected. The Dutch syndromes were designated as Oppositional, Withdrawn/Depressed, Aggressive, Anxious, Overactive, Sleep Problems, and Somatic Problems. CFA indicated that this seven-syndrome model fit the data in each sample. The Dutch syndrome model was similar to the U.S. CBCL/2–3 syndrome model, with all Dutch syndromes having U.S. counterparts, except the Overactive syndrome. However, the use of different factor analytic procedures makes it difficult to compare the Koot and Achenbach CBCL/2–3 results. The authors concluded that they were "unable to decide to what extent the somewhat different factor structure of the CBCL/2–3 obtained in this study compared to the one found for American samples was the result of the use of different factor-analytic methods and different samples, or was a reflection of true cross-cultural differences in young preschoolers' problem behaviors" (p. 194).[9]

Factor analytic studies of the preschool CBCL have thus produced somewhat mixed results. The CBCL/1.5–5 syndrome model fit the data obtained for adopted Chinese girls,[8] and the Dutch CBCL/2–3 model was similar to the U.S. CBCL/2–3 model.[9] Although the Konold et al. results should be interpreted with caution, they indicated better fit when the Emotionally Reactive and Anxious/Depressed factors were combined into a single syndrome. However, this six-syndrome model had limited invariance across ethnic groups and SES.

The foregoing studies tested preschool syndromes in different ways in limited populations. To evaluate the wider generalizability of syndromes, it is important to uniformly test them in multiple populations. We refer to the populations included in the present study as "societies" rather than as "countries" or "cultures," because not all were countries (e.g., Flanders in Belgium), and there were cultural variations within as well as between them.

Study Purpose

The purpose of this study was to test the configural invariance of the correlated seven-syndrome model of the CBCL/1.5–5[4] in 23 societies. We believe this study to be the first multicultural test of taxonomic constructs of preschool psychopathology. It differs from other studies of preschool psychopathology by using uniform CFA procedures to test the same syndrome model in 23 societies. If a syndrome model is supported in many societies, it can provide a taxonomy for assessing psychopathology in those societies. Such a taxonomy can facilitate international communication and collaboration in research, training, and clinical care. It can also lay foundations for taxonomic research to take account of differences found between societies and to pinpoint possible reasons for the differences.

METHOD

Samples

We analyzed data for 19,106 children 1.5 to 5 years old from the 23 epidemiological samples listed in Table 1. We included children who were referred for mental health services. The English language CBCL/1.5–5 was translated for use in all societies except in Australia. Independent back-translations established that translations captured the original meanings. Consent requirements for each investigator's institution were fulfilled. All samples were approximately 50% male. In a companion paper comparing scale scores for our samples, plus the U.S. sample, Rescorla et al. (unpublished data, August 2010) present details of how the samples were obtained. Consistent with standard procedures for scoring the CBCL/1.5–5[4], we excluded forms with more than eight missing item ratings (2% of forms in two societies; ≤1% in 21 societies). For forms with eight or fewer missing items, missing items were scored as zero.

To determine whether any of the 67 items in the syndrome model were selectively omitted by many parents, we tabulated the percentage of parents who failed to rate each item in each of the five societies having the largest percentage of omitted items (China, Finland, Lithuania, Singapore, and Iceland). The largest percentage of omissions was 4% for item *15. Defiant* in the Singapore sample and also for item *24. Doesn't eat well* in the China sample. The second largest was

TABLE 1 Results of Confirmatory Factor Analyses (CFA) for 23 Societies

Society	N	RMSEA	CFI	TLI	Item Loadings	
					Mean (SD)	Median
1. Australia	1,793	0.037	0.878	0.921	0.60 (0.11)	0.62
2. Belgium (Flanders)	1,117	0.048	0.885	0.921	0.64 (0.10)	0.65
3. Chile	483	0.047	0.904	0.930	0.61 (0.12)	0.63
4. China	908	0.038	0.921	0.963	0.63 (0.10)	0.63
5. Denmark	851	0.040	0.842	0.928	0.67 (0.10)	0.68
6. Finland	370	0.048	0.904	0.926	0.65 (0.15)	0.69
7. France	1,204	0.044	0.859	0.908	0.61 (0.09)	0.62
8. Germany	850	0.045	0.882	0.934	0.63 (0.09)	0.65
9. Iceland	305	0.054	0.840	0.862	0.59 (0.18)	0.60
10. Iran	1,346	0.053	0.789	0.891	0.60 (0.13)	0.62
11. Italy	466	0.050	0.799	0.824	0.54 (0.14)	0.55
12. Korea	2,327	0.046	0.847	0.933	0.64 (0.10)	0.66
13. Kosovo	481	0.041	0.929	0.942	0.61 (0.12)	0.63
14. Lithuania	931	0.045	0.853	0.901	0.57 (0.13)	0.58
15. Netherlands	608	0.046	0.875	0.901	0.60 (0.12)	0.62
16. Peru	1,027	0.036	0.885	0.920	0.56 (0.10)	0.58
17. Portugal	407	0.045	0.878	0.907	0.60 (0.11)	0.61
18. Romania	938	0.059	0.825	0.893	0.62 (0.13)	0.63
19. Singapore	301	0.044	0.952	0.972	0.68 (0.11)	0.69
20. Spain	570	0.044	0.878	0.896	0.59 (0.15)	0.61
21. Taiwan	306	0.051	0.926	0.950	0.67 (0.12)	0.67
22. Turkey	825	0.043	0.876	0.912	0.55 (0.10)	0.56
23. United Arab Emirates	692	0.048	0.862	0.904	0.61 (0.14)	0.62

Note: For Finland, Turkey, and United Arab Emirates, the CFAs were performed on imputed data sets. CFI = comparative fit index; RMSEA = root mean square error of approximation; TLI = Tucker-Lewis Index.

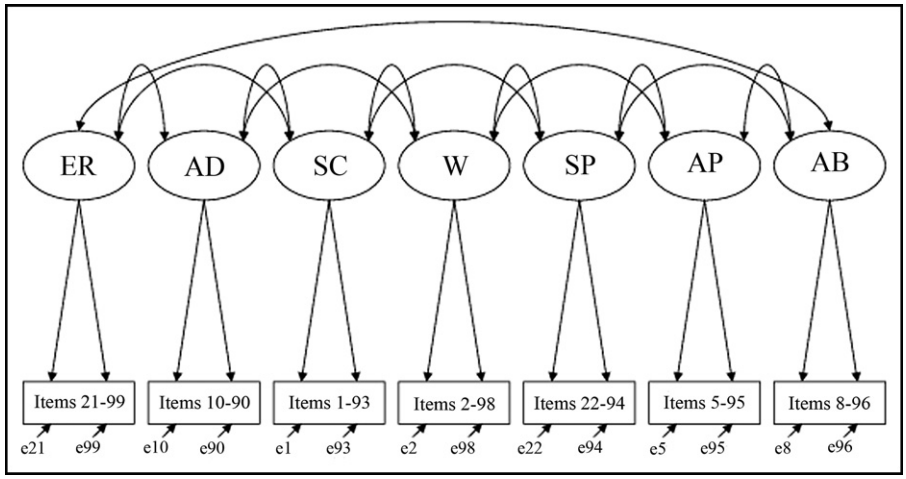
3% for item *74. Sleeps less than most children* in the China sample. As missing ratings were thus rare and did not appear to be caused by significant cultural differences in item meanings, they were considered unlikely to affect the CFA results and did not justify imputation of missing data.

The CBCL/1.5–5[4] was used for all samples except Finland, United Arab Emirates (UAE), and 685 of the 885 Turkish children, where the CBCL/2–3[7] was used. As detailed by Achenbach and Rescorla,[4] two CBCL/2–3 items were replaced on the CBCL/1.5–5 with two new items: *51. Shows panic for no good reason* and *79. Rapid shifts between sadness and excitement*. Both items load on the Emotionally Reactive syndrome. To take advantage of data collected with the CBCL/2–3 in testing the CBCL/1.5–5 syndrome model, multiple imputation was used for items 51 and 79 in samples assessed with the CBCL/2–3. Multiple imputation was chosen over single imputation because of its effective handling of error variance.[10] For each sample assessed with the CBCL/2–3, 10 multiply imputed data sets were generated using the imputation by chained equations module of STATA.[11] This module applies a prediction equation specified by the user for a reference population with complete data using multiple imputation by chained equations.[12,13] An ordinal logistic regression predicting the missing 0-1-2 values from the remaining seven items of the Emotionally Reactive syndrome served as the prediction equation. For Turkey, the Turkish subsample collected with the CBCL/1.5–5 served as the reference population, whereas for UAE and Finland, the CBCL/1.5–5 U.S. normative sample was used as the reference population.[4] The U.S. normative sample was selected as the reference population for the prediction equation on which to base the imputed values for items 51 and 79 in the Finnish and UAE samples because their sample means on the Total Problems score (the sum of all items) fell within 0.1 and 0.02 SD units, respectively, of the Total Problems score for the U.S. normative sample. For Turkey, UAE, and Finland, CFAs were conducted on the ten multiply imputed data sets simultaneously via Mplus.[14]

Tested Model

We tested the correlated seven-factor model presented in Figure 1. Each item was assigned to only one factor. To identify the metric for a factor, we set the loading of one item to 1.0.

FIGURE 1 For the purpose of clear presentation, some latent factor correlations are not depicted. Note: AB = aggressive behavior; AD = anxious/depressed; AP = attention problems; ER = emotionally reactive; S = sleep; SC = somatic complaints; W = withdrawn.



Data Analyses

Parents rated CBCL items as 0 = not true (as far as you know), 1 = somewhat or sometimes true, and 2 = very true or often true, based on the preceding 2 months. To account for nonnormal distributions, we used the weighted least squares with standard errors and mean- and variance-adjusted χ^2 estimator (WLSMV) implemented via Mplus 5.1.[14] Items were dichotomized as 0 versus 1 or 2 for the computation of tetrachoric correlations. The root mean square error of approximation (RMSEA) served as the primary fit index, because it performed most robustly in a Monte Carlo simulation study that tested model fit indices with binary categorical data across model specification and complexity conditions.[15] An RMSEA cutoff of 0.05 has been proposed to indicate good fit[15,16] and a cutoff of 0.08 to indicate acceptable fit.[16] We also computed the Comparative Fit Index (CFI)[17] and the Tucker-Lewis Index (TLI),[18] but regarded them as secondary because they did not perform as well as the RMSEA across the simulation conditions.[15] Hu and Bentler[19] proposed CFI and TLI values >0.95 for good model fit. We used criteria of >0.90 for good fit and 0.80 to 0.90 for acceptable fit, as recommended by Browne and Cudeck,[16] because the Hu and Bentler[19] criterion has been criticized for rejecting correctly specified complex models.[20]

RESULTS

The model converged for all samples. As presented in Table 1, RMSEAs ranged from 0.036 to 0.059, indicating acceptable to good fit for all societies (25^{th} percentile = 0.043, 50^{th} = 0.045, and 75^{th} = 0.048). CFIs ranged from 0.789 to 0.952, indicating acceptable to good fit for all societies except Iran and Italy (25^{th} percentile = 0.847, 50^{th} = 0.878, and 75^{th} = 0.904). TLIs ranged from 0.824 to 0.972 (25^{th} percentile = 0.901, 50^{th} = 0.920, and 75^{th} = 0.933), indicating acceptable to good fit for all societies.

For 14 societies, the model converged smoothly. For nine societies (China, Iceland, Iran, Italy, Kosovo, the Netherlands, Portugal, Romania, and Singapore), the latent variable variance/covariance matrix was non-positive definite. A symmetrical matrix such as the latent variable variance/covariance matrix is considered to be positive definite when the matrix (including its principal submatrices) has a positive determinant. Causes of nonpositive definiteness include out-of-range parameters (i.e., a negative variance or residual variance for a latent variable, or correlations exceeding or equaling 1.0 between two latent variables) or a linear dependency among more than two latent variables.[14]

For Iran and Kosovo, the correlation between the Emotionally Reactive and Anxious/Depressed latent variables slightly exceeded 1.00 (1.009 and 1.003, respectively). Thus, only these two (0.04%) of the 5,267 estimated parameters were out-of-range, causing the nonpositive definite matrix. To test the two out-of-range parameters, we used the technique of Van Driel[21] recommended by Chen et al.[22] and McDonald.[23] It involves forming confidence intervals around the out-of-range parameter, and determining whether the confidence interval overlaps with the admissible param-

eter space. If there is overlap, then the out-of-range parameter estimate could be attributable to sampling fluctuations, rather than to model specification error. Van Driel's[21] technique showed that the 95% confidence intervals around both out-of-range parameters contained admissible values.

For the remaining seven societies (China, Iceland, Italy, the Netherlands, Portugal, Romania, and Singapore), the correlations between the Emotionally Reactive and Anxious/Depressed latent variables were 0.995, 0.886, 0.981, 0.884, 0.983, 0.947, and 0.999, respectively. Given the high correlations in the seven societies and the two out-of-range parameters for Iran and Kosovo, we combined the Emotionally Reactive and Anxious/Depressed factors and tested a six-factor model for the nine societies. The six-factor model converged smoothly for all nine societies. The fit of the six-factor model was compared to the fit of the seven-factor model using the Mplus DIFFTEST option, which is based on the χ^2 difference test developed for use with the WLSMV estimator.[14] For seven of the nine societies (China, Iceland, Italy, Iran, the Netherlands, Portugal, and Romania), the results of the DIFFTEST indicated that combining the Emotionally Reactive and Anxious/Depressed factors led to a significant worsening of model fit. For Kosovo and Singapore, the fit of the six- and seven-factor models was not statistically different.

For 17 societies, all 67 items loaded significantly on their predicted factors. For Finland, Kosovo, Lithuania, and UAE, one item failed to reach statistical significance (items 7, 19, 7, and 86, respectively). For Italy and Spain, two items (items 19 and 56 for Italy, and 46 and 95 for Spain), and for Iceland, four items (items 12, 46, 67, and 86) failed to reach statistical significance. Table 1 presents descriptive statistics for item loadings for each society, including the mean, median, and standard deviation. Mean factor loadings for each society ranged from 0.54 (Italy) to 0.68 (Singapore). The mean of the mean factor loadings (i.e., averaged across all societies) was 0.61 (25^{th} percentile = 0.59, 50^{th} = 0.61, 75^{th} = 0.64). Table 2 presents descriptive statistics for factor loadings for each item and syndrome across societies, including the mean, median, and standard deviation. Mean factor loadings across societies ranged from 0.43 (item *7. Can't stand having things out of place*) to 0.76 (item *82. Sudden changes in mood or feelings)*. The mean of the mean factor loadings across societies was .61 (25^{th} percentile = 0.56, 50^{th} = 0.62, 75^{th} = 0.68). Finally, mean factor loadings by syndrome ranged from 0.55 (Somatic Complaints) to 0.64 (Attention Problems and Aggressive Behavior).

DISCUSSION

We found that the seven-syndrome model fit the data for all 23 societies. Our primary fit index, the RMSEA, indicated acceptable to good fit in all societies, although the CFI and TLI were more variable. Of the 5,267 estimated parameters, the two (0.04%) that were outside the allowable parameter space possibly reflected sampling error, according to Van Driel.[21] Across societies, the median loading of items on their respective factors was high (0.61).

The findings supported the configural invariance of the seven-syndrome structure in the tested societies. This indicates that the seven syndromes capture patterns of young children's emotional, behavioral, and social problems manifested in ratings by a great variety of parents. The fit between the seven-syndrome structure and data from 23 societies thus supports the taxonomic generalizability of the seven-syndrome structure. Clinicians, researchers, and educators working with preschoolers from the 23 societies can therefore assess, describe, and communicate about their parent-rated behavioral, emotional, and social problems in terms of the seven syndromes.

Our samples came from Asia, Australasia, Europe, the Middle East, and South America. The societies vary greatly in characteristics affecting views of children and childrearing practices, including political, social, and economic systems, ethnicity, language, and religion. Differences in sample characteristics (e.g., response rate) and methodological variations (e.g., recruitment procedures) could also have affected patterns of problems. Despite these differences, the seven syndromes were found to reflect similar patterns of parent-rated problems in all the societies. Albeit with important developmental differences in the syndromes, Ivanova et al.[24] also found support in 30 societies for the eight syndromes scored from the CBCL for Ages 6–18.[25]

Others have cited resistance to recognizing that very young children experience significant psychological problems and that these problems need to be carefully studied in their own right.[26,27] Our findings support efforts to systematize analyses of preschoolers' emotional, behav-

TABLE 2 Mean and Median Item Loadings for Child Behavior Checklist for Ages 1.5–5 Syndromes Across 23 Societies

Syndrome Item	Mean Loading (SD)	Median Loading
Emotionally reactive	**0.60 (0.13)**	**0.62**
21. Disturbed by any change in routine	0.57 (0.08)	0.57
46. Nervous movements or twitching	0.45 (0.17)	0.51
51. Shows panic for no good reason	0.56 (0.13)	0.60
79. Rapid shifts between sadness and excitement	0.68 (0.10)	0.69
82. Sudden changes in mood or feelings	0.76 (0.06)	0.75
83. Sulks a lot	0.66 (0.07)	0.64
92. Upset by new people or situations	0.51 (0.08)	0.53
97. Whining	0.66 (0.05)	0.66
99. Worries	0.56 (0.10)	0.53
Anxious/depressed	**0.61 (0.11)**	**0.62**
10. Clings to adults or too dependent	0.54 (0.10)	0.55
33. Feelings are easily hurt	0.57 (0.08)	0.60
37. Gets too upset when separated from parents	0.54 (0.07)	0.54
43. Looks unhappy without good reason	0.69 (0.07)	0.68
47. Nervous, highstrung, or tense	0.71 (0.08)	0.70
68. Self-conscious or easily embarrassed	0.51 (0.08)	0.50
87. Too fearful or anxious	0.63 (0.10)	0.64
90. Unhappy, sad, depressed	0.70 (0.11)	0.72
Somatic complaints	**0.55 (0.16)**	**0.57**
1. Aches or pains (without medical cause)	0.57 (0.12)	0.59
7. Can't stand having things out of place	0.43 (0.17)	0.47
12. Constipated, doesn't move bowels (when not sick)	0.49 (0.15)	0.51
19. Diarrhea or loose bowels (when not sick)	0.49 (0.17)	0.51
24. Doesn't eat well	0.49 (0.08)	0.50
39. Headaches (without medical cause)	0.60 (0.14)	0.65
45. Nausea, feels sick (without medical cause)	0.70 (0.12)	0.71
52. Painful bowel movements (without medical cause)	0.61 (0.15)	0.63
78. Stomachaches or cramps (without medical cause)	0.65 (0.11)	0.63
86. Too concerned with neatness or cleanliness	0.46 (0.19)	0.50
93. Vomiting, throwing up (without medical cause)	0.57 (0.13)	0.60
Withdrawn	**0.61 (0.12)**	**0.62**
2. Acts too young for age	0.49 (0.08)	0.49
4. Avoids looking others in the eye	0.53 (0.11)	0.52
23. Doesn't answer when people talk to him/her	0.64 (0.07)	0.66
62. Refuses to play active games	0.60 (0.10)	0.58
67. Seems unresponsive to affection	0.62 (0.16)	0.64
70. Shows little affection toward people	0.70 (0.08)	0.70
71. Shows little interest in things around him/her	0.65 (0.10)	0.65
98. Withdrawn, doesn't get involved with others	0.65 (0.11)	0.69
Sleep problems	**0.62 (0.09)**	**0.62**
22. Doesn't want to sleep alone	0.51 (0.07)	0.50
38. Has trouble getting to sleep	0.68 (0.08)	0.66
48. Nightmares	0.65 (0.07)	0.65
64. Resists going to bed at night	0.68 (0.07)	0.69
74. Sleeps less than most kids during day and/or night	0.60 (0.08)	0.59
84. Talks or cries out in sleep	0.60 (0.08)	0.60
94. Wakes up often at night	0.65 (0.06)	0.66
Attention problems	**0.64 (0.14)**	**0.64**
5. Can't concentrate, can't pay attention for long	0.66 (0.10)	0.66
6. Can't sit still, restless, or hyperactive	0.69 (0.09)	0.69
56. Poorly coordinated or clumsy	0.57 (0.13)	0.58
59. Quickly shifts from one activity to another	0.72 (0.10)	0.74

TABLE 2 Continued

Syndrome Item	Mean Loading (SD)	Median Loading
95. Wanders away	0.54 (0.16)	0.57
Aggressive behavior	**0.64 (0.09)**	**0.65**
8. Can't stand waiting: wants everything now	0.66 (0.06)	0.68
15. Defiant	0.66 (0.08)	0.65
16. Demands must be met immediately	0.69 (0.09)	0.70
18. Destroys things belonging to his/her family or other children	0.56 (0.07)	0.54
20. Disobedient	0.71 (0.07)	0.70
27. Doesn't seem to feel guilty after misbehaving	0.56 (0.07)	0.55
29. Easily frustrated	0.62 (0.09)	0.64
35. Gets in many fights	0.60 (0.10)	0.61
40. Hits others	0.61 (0.08)	0.62
42. Hurts animals or people without meaning to	0.56 (0.06)	0.54
44. Angry moods	0.71 (0.07)	0.71
53. Physically attacks people	0.66 (0.06)	0.66
58. Punishment doesn't change his/her behavior	0.62 (0.09)	0.62
66. Screams a lot	0.68 (0.08)	0.68
69. Selfish or won't share	0.56 (0.07)	0.56
81. Stubborn, sullen, or irritable	0.72 (0.07)	0.74
85. Temper tantrums or hot temper	0.71 (0.08)	0.67
88. Uncooperative	0.64 (0.11)	0.63
96. Wants a lot of attention	0.66 (0.07)	0.66

Note: Values in boldface type are descriptive statistics for entire syndromes.

ioral, and social problems.[2] We are currently testing whether problems reported by preschool teachers and daycare providers manifest similar patterns across diverse societies.

The findings of similar patterns of problems across 23 societies indicate that we can meaningfully compare scores on the seven tested syndromes among these societies. In a companion study, Rescorla et al. (unpublished data, August 2010) compared CBCL/1.5–5 syndrome and *DSM*-oriented scale scores for children in the 23 societies, plus a U.S. general population sample. The effect sizes for differences among societies were small to medium.[28] Although modest in magnitude, the significant effects of society on scale scores revealed potentially important variations in levels of problems reported by parents in different societies.

Canino and Alegria[29] have contrasted studies that, like ours, test syndromes in various populations versus studies that, like Rescorla's—test *differences* between levels of problems. As an example, a particular attention-deficit/hyperactivity disorder (ADHD) syndrome model could be applicable in various populations, whereas the populations differ in the prevalence of ADHD defined by that model.

Canino and Alegria also contrasted hypotheses that syndromes are universal versus relative to particular populations versus mixtures of universal and population-relative aspects. These contrasting hypotheses can be tested by determining whether particular syndromes are supported in different populations. Because the universal hypothesis is constrained by the impracticality of testing syndromes in all human populations, we prefer the term "culture-general" to "universal." Testing particular syndromes in different populations requires etic[30] research using the same standardized assessment in all the populations. By contrast, testing of population-relative characteristics requires emic[30] assessment tailored to particular populations. Our *etic* findings support particular syndromes in many populations. However, our finding that Emotionally Reactive versus Anxious/Depressed syndromes may be less distinguishable in some societies than in others invites emic research on possible reasons for this difference.

The large and diverse samples and the uniformity of analytic methods across 23 societies are strengths of this study, but certain limitations should also be noted. First, the nonnormal distribution of our data required use of the WLMSV estimator, a method so computationally intensive

that only configural invariance of our complex model could be tested. A second limitation is that the seven tested syndromes do not necessarily include all possible syndromes. Although CFA tests the fit of a specific model, other syndrome models might also fit the data. For example, we found that for nine societies a six-factor model that combined the Emotionally Reactive and Anxious/Depressed syndromes also fit the data. However, this model fit significantly worse than the seven-factor model in seven of the nine societies. In addition, in the absence of reliable power analytic procedures for CFAs of complex models such as ours using the WLSMV, it is difficult to ascertain whether sample sizes as small as 301, 305, and 306 (Singapore, Iceland, and Taiwan) might have affected the results, although the seven-syndrome model fit these samples as well as it fit several larger samples. Finally, because we could not analyze samples from all human societies, our findings do not reflect all possible variations in model fit across all societies.

Mental health clinicians around the world are increasingly called upon to assist children of diverse backgrounds. For example, in the United States, the percentage of children with at least one foreign-born parent rose from 15% in 1994 to 22% in 2008.[31] Findings from this study and from the companion Rescorla et al. study indicate that clinicians can use the seven syndromes as foci for assessing, describing, and communicating about problems reported by parents for preschoolers from the 23 societies. The seven syndromes can be used at multiple stages of clinical services, from assessment and case formulation to treatment design and outcome evaluation. The syndromes can also be applied to many children who cross cultures, including immigrants, refugees, international adoptees, children attending international schools, and children in families of mixed heritages. To help clinicians take account of societal differences in the distributions of scale scores, the Rescorla et al. findings for the preschool syndromes and *DSM*-oriented scales have been incorporated into multicultural norms.[32]

The growing awareness of the mental health needs of young children, coupled with the increasing interconnectedness of societies, highlights needs for psychometrically sound tools for assessing preschoolers' problems in diverse societies. The seven syndromes tested in the present study can be easily assessed in diverse societies using parents' reports. The syndromes offer mental health professionals a taxonomic framework supported by data from many societies and practical in its operationalization and clinical use. This framework also provides a basis for further research on societal differences in the patterning of problems and on possible reasons for the differences. &

Accepted September 28, 2010.

Drs. Ivanova, Achenbach, and Harder are with the University of Vermont. Dr. Rescorla is with Bryn Mawr College. Dr. Ang is with Nanyang Technological University. Drs. Bilenberg and Ms. Kristensen are with the University of Southern Denmark. Dr. Bjarnadottir is with Glaesibaer Health Clinic. Drs. Capron and Duyme are with the University of Paris. Ms. De Pauw is with Ghent University. Dr. Dias is with Portugese Catholic University. Dr. Dobrean is with Babes-Bolyai University. Drs. Doepfner and Plueck are with the University of Cologne. Dr. Eapen is with the University of New South Wales. Dr. Erol is with Ankara University. Dr. Esmaeili is with the Institute for Exceptional Children. Dr. Ezpeleta is with the Universitat Autonoma de Barcelona. Dr. Frigerio and Ms. Montirosso are with the Scientific Institute E. Medea. Dr. Gonçalves is with the University of Minho. Mr. Gudmundsson is with the University of Iceland. Dr. Jeng and Ms. Wu are with National Taiwan University. Ms. Jetishi is with the University of Dardania. Dr. Jusiene is with Vilnius University. Dr. Kim is with Hunoconsulting. Mr. Lecannelier is with Universidad del Desarrollo. Dr. Leung is with the Chinese University of Hong Kong. Dr. Liu is with the University of Pennsylvania. Dr. Oh is with Yonsei University. Drs. Pomalima and Valverde are with Peruvian National Institute of Mental Health. Dr. Shahini is with the University Clinical Center of Kosova. Dr. Silva is with the Universidad de La Frontera. Dr. Simsek is with Harran University. Dr. Sourander is with Turku University and Turku University Hospital. Dr. Van Leeuwen is with Leuven University. Ms. Woo is with the Child Guidance Clinic of the Singapore Institute of Mental Health. Dr. Zubrick is with the Curtin Centre for Developmental Research, Curtin University of Technology. Dr. Verhulst is with Erasmus University Medical Center-Sophia's Children's Hospital.

Disclosure: Dr. Ivanova receives research and salary support from the Research Center for Children, Youth, and Families, which publishes the Child Behavior Checklist (CBCL). Dr. Achenbach is President of the Research Center for Children, Youth, and Families, and receives remuneration. Dr. Rescorla receives remuneration from the Research Center for Children, Youth, and Families. Dr. Harder previously held a University of Vermont Postdoctoral Fellowship funding by the Research Center for Children, Youth, and Families. Drs. Bjarnadottir, Gudmundsson, Leung, Verhulst, and Mr. Gudmundsson, receive research support from the sale of the CBCL. Dr. Bilenberg has received honoraria from Eli Lilly and Co., Novartis, Neuroscience, and Janseen Cilag. He has received research support from the Danish Research Foundation, the Lundbeck Foundation, Hermansens Mindelegat, and Mads Clausen Fond. Dr. Eapen has received research support from the Sheikh Hamdan Bin Rashid Al Maktoum Award for Medical Sciences. Dr. Jusiene has received research support from the Lithuanian Science and Studies Foundation. Drs. Ang, Capron, Dias, Dobrean, Doepfner, Duyme, Erol, Esmaeili, Ezpeleta, Frigerio, Gonçalves, Jeng, Kim, Liu, Oh, Plueck, Pomalima, Shahini, Silva, Simsek, Sourander, Valverde, Van Leeuwen, and Zubrick, Ms. De Pauw, Ms. Kristensen, Mr. Lecannelier, Ms. Montirosso, Ms. Jetishi, Ms. Woo, and Ms. Wu report no biomedical financial interests or potential conflicts of interest.

Correspondence to: Masha Y. Ivanova, Ph.D., Vermont Center for Children, Youth and Families, University of Vermont, 1 South Prospect Street, Burlington, VT 05401. e-mail: Masha.Ivanova@uvm.edu

0890-8567/$36.00/©2010 American Academy of Child and Adolescent Psychiatry

DOI: 10.1016/j.jaac.2010.08.019

REFERENCES

1. Egger HL. Psychiatric assessment in young children. Child Adolesc Psychiatr Clin North Am. 2009;18:559-580.
2. Egger HL, Angold A. Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. J Child Psychol Psychiatry. 2006;47:313-337.
3. Postert C, Averbeck-Holocher M, Beyer T, Müller J, Furniss T. Five systems of psychiatric classification for preschool children: Do differences in validity, usefulness, and reliability make for competitive or complimentary constellations? Child Psychiatry Hum Dev. 2009;40:25-41.
4. Achenbach TM, Rescorla LA. Manual for the ASEBA Preschool Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families; 2000.
5. Vandenberg RJ, Lance CE. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. Organ Res Methods. 2000;3:4-70.
6. Konold TR, Hamre BK, Pianta R. Measuring problem behaviors in young children. Behav Disord. 2003;28:111-123.
7. Achenbach TM. Manual for the Child Behavior Checklist/2-3 and 1992 Profile. Burlington, VT: University of Vermont Department of Psychiatry; 1992.
8. Tan TX, Dedrick RF, Marfo K. Factor structure and clinical implications of Child Behavior Checklist/1.5–5 ratings in a sample of girls adopted from China. J Pediatr Psychol. 2007;32:807-818.
9. Koot HM, Van den Oord EJCG, Verhulst FC, Boomsma DL. Behavioral and emotional problems in young preschoolers: Cross-cultural testing of the validity of the Child Behavior Checklist/2-3. J Abnorm Child Psychol. 1997;25:183-196.
10. Graham J. Missing data analysis: Making it work in the real world. Annu Rev Psychol. 2009;60:549-576.
11. Stata Corporation. Stata Statistical Software: Release 10. College Station, TX: Stata Corporation; 2007.
12. Raghunathan TE, Lepkowski JM, van Hoewyk J, Solenberger PA. Multivariate technique for multiply imputing missing values using a sequence of regression models. Surv Methodol. 2001;27:85-95.
13. Van Buuren S, Brand JPL, Groothius-Oudshoorn, CGM, Rubin DB. Fully conditional specification in multivariate imputation. J Stat Comput Simulation. 2006;76:1049-1064.
14. Muthén LK, Muthén BO. Mplus User's guide. 5th ed. Los Angeles, CA: Muthén & Muthén; 2007.
15. Yu CY, Muthén BO. Evaluation of Model Fit Indices for Latent Variable Models with Categorical and Continuous Outcomes (Technical Report). Los Angeles, CA: UCLA, Graduate School of Education and Information Studies; 2002.
16. Browne MW, Cudeck R. Alternative ways of assessing model fit. In Bollen KA, Long JS, eds. Testing Structural Equation Models. Newbury Park, CA: Sage; 1993:136-162.
17. Bentler PM. Comparative fit indexes in structural models. Psychol Bull. 1990;107:238-246.
18. Tucker LR, Lewis C. The reliability coefficient for maximum likelihood factor analysis. Psychometrika. 1973;38:1-10.
19. Hu L, Bentler P. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equation Modeling. 1999;6:1-55.
20. Marsh HW, Hau KT, Wen Z. In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. Struct Equation Modeling. 2004;11:320-341.
21. Van Driel, OP. On various causes of improper solutions in maximum likelihood factor analysis. Psychometrika. 1978;43:225-243.
22. Chen F, Bollen KA, Paxton P, Curran PJ, Kirby JB. Improper solutions in structural equation models: Causes, consequences, and strategies. Sociol Methods Res. 2001;25:223-251.
23. McDonald RP. Respecifying improper structures. Struct Equation Modeling. 2004;11:194-209.
24. Ivanova MY, Achenbach TM, Dumenci L, Rescorla LA, Almqvist F, Bilenberg N, *et al*. Testing the 8-syndrome structure of the Child Behavior Checklist in 30 societies. J Clin Child Adolesc Psychol. 2007;36;405-417.
25. Achenbach TM, Rescorla LA. Manual for the ASEBA School-Age Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families; 2001.
26. Angold A, Egger HL. Psychiatric diagnosis in preschool children. In DelCarmen-Wiggins R, Carter A, eds. Handbook of Infant, Toddler, and Preschool Mental Health Assessment. New York, NY: Oxford University Press; 2004:123–139.
27. Carter AS, Briggs-Gowan MJ, Davis NO. Assessment of young children's social-emotional development and psychopathology: Recent advances and recommendations for practice. J Child Psychol Psychiatry. 2004;45:109-134.
28. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. New York, NY: Academic Press; 1988.
29. Canino G, Alegria, M. Psychiatric diagnosis—is it universal or relative to culture? J Child Psychol Psychiatry. 2008;49:237-250.
30. Pike KL. Language in Relation to a Unified Theory of the Structure of Human Behavior. The Hague: Mouton & Co: 1967.
31. Federal Interagency Forum on Child and Family Statistics. America's Children: Key National Indicators of Well-Being, 2009. Washington, DC: U.S. Government Printing Office; 2009.
32. Achenbach TM, Rescorla LA. Multicultural Supplement to the Manual for the ASEBA Preschool Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families; 2010.